Remarks of Dr. Daniel A. Long
Assistant Professor of Sociology at Wesleyan University
Middletown, CT

Before the Education Committee
On Senate Bill 24, Section 28 and 30
Topic: Teacher Evaluation

February 21, 2012

Good afternoon Senator Stillman and Representative Fleischman, and members of the Education Committee.

My name is Dr. Daniel Long, I am a Sociology of Education professor from Wesleyan University and I have studied the causes of the achievement gap and value added growth models for the last 15 years. I am here today to comment on Senate Bill 24, Section 28 and 30.

I have reviewed the best research on the achievement gap and teacher evaluation models and I can say with certainty that the current proposed educational reforms will do almost nothing to narrow the achievement gap and will lower the average achievement in CT for all students.

Using standardized tests to evaluate teachers is ineffective, unfair and counterproductive.

First, the use of these tests is inaccurate and ineffective

In my own study of the performance of more than 200,000 students in school systems in 64 countries, I found that using student test scores to evaluate teachers tends to lower overall student learning outcomes[i]. This occurs because teachers are given an incentive to teach to the narrow test questions instead of promoting an in-depth knowledge of academic subjects and critical thinking skills.

Second, the use of tests is unfair.

The use of student tests to evaluate teachers also blames teachers for factors outside of their control. A students' academic growth is due to in large part to parents' education, family support, parents' resources, and peer effects. Therefore the value added scores in teacher evaluations will reflect the vast inequality in CT towns and cities.

Third, the use of tests is unreliable, untested, and counterproductive.

Other than my own study that I mention above, there have been almost no studies of the long term effects of actual test evaluation programs. All claims about the benefits of using student tests to evaluate teachers are based on hypothetical statistical calculations that fail to consider

the negative effects of teaching to the test, the narrowing of the curriculum, the decline in teacher cooperation, and the decline in teacher moral.

At a minimum the state needs to run pilot programs for several years in rich and poor towns in Connecticut.

This reform is like proscribing a new drug nationwide without field trials. The medical community would never allow this. And we should not allow this state wide implementation without a multiyear state wide evaluation program first.

In sum, using standardized tests to evaluate teachers is ineffective, unfair, and counterproductive.

The proposed use of standardized tests scores in teacher evaluation at best will increase teaching to the test, decrease overall student knowledge and critical thinking skills, and create an unfair evaluation system that rewards teachers from wealthy towns and punishes teachers in poor towns and cities. At worst this system lower teacher quality, lower teacher morale, and cause the state of CT go from one of the top performing states in average performance to a low performing state.

You might ask, if using student tests are not the answer what can be done? There are a wide variety of exceptional qualitative teacher evaluation programs and professional development programs. In the 1990s Connecticut was a model for the country in teacher evaluation using highly trained professional evaluators. This plus an expansion of professional development helped Connecticut become one of the highest achieving states in the country.

Unfortunately, the funding for this excellent program was cut.

Instead of the current teacher evaluation proposals in this bill we should return to a fully funded model of highly trained professional evaluators.

Also, the main problem with education in Connecticut is not teachers. Teachers in the state are among the best nationwide, due in large part to the state's rigorous credentialing standards.

The problem with education in Connecticut is income inequality, not teacher quality. Unfortunately, the plans Gov. Dannel P. Malloy has outlined for education reform — for the most part — take us in entirely the wrong direction.

Education in Connecticut is a paradox. Though the National Assessment of Educational Progress consistently ranks the state among the highest scoring for student achievement, we also suffer from the highest black/white and poor/non-poor achievement gaps in the country.

For example, look at 2011 Connecticut Mastery Test data for eighth-graders. There is a dramatic difference in the percent of students testing at goal in mathematics in the wealthy

towns of Westport, Darien and Ridgefield — 96, 94 and 92 percent, respectively — as compared with less wealthy Waterbury, New Britain and New London — 28.5, 21 and 19 percent, respectively. Almost every town in Connecticut follows this strong correlation between income and achievement scores.

If schools were primarily responsible for the achievement gap, we would expect the gap to increase as students progressed through grade school. Yet the gap is almost unchanged from third to eighth grade, as evidenced by results of the 2011 CMTs. The inequalities begin before students enter school.

The governor claims that the source of the achievement gap in Connecticut is teacher quality and the lack of school choice. He proposes the following solutions: First, use student tests to evaluate teachers; second, weaken the state's standards for teacher credentials; third, increase school choice; fourth, provide a small increase in funding for pre-kindergarten instruction; and fifth, minimal funds to narrow the funding gap for only a few schools . Research shows that this reform plan is wrong on the first three counts, and only partially correct on the fourth and fifth points.

First, using standardized tests to evaluate teachers is ineffective, unfair and counterproductive as I discussed above.

Second, allowing individual school districts to modify the requirements to become a teacher will result in lower teacher quality. This is especially true in the lowest-performing school districts, which will likely feel pressure to lower their standards in order to attract teachers to low-performing urban public schools.

Third, the best study of school choice nationwide, Stanford University's Center for Research on Education Outcomes, matched students by race/ethnicity and socioeconomic status and found that charter schools are equal or worse than public schools in academic achievement. While there are a handful of exceptional charter schools, on average charter schools are no better than public schools and should not be at the center of an educational reform policy.

On the fourth point, Malloy's emphasis on high-quality preschool — which experimental evidence has shown to dramatically increase student achievement — is a step in the right direction. However, this component of the plan is underfunded. In a state with 204,000 children under the age of 5, the governor's plan only provides for an additional 500 preschool slots. Preschool programs and community-level early interventions, such as those pioneered by the Harlem Children's Zone, should be made an even greater priority in Connecticut.

Fifth, the changes in the funding formula will only partially help a handful of the poorest cities with an increase in $40 million dollars in funding. These funds are inadequate to address the existing inequalities in funding. In reality, the increased funding needed is closer to $600 million not $40 million dollars.

What else can be done? The biggest impact would come from policies that equalize the average socioeconomic status in all schools. Such economic integration can only be achieved by expanding requirements in the 1996 Connecticut Supreme Court case, Sheff v. O'Neill, to address economic integration in addition to racial integration. A politically difficult, but ideal, solution would be to merge adjacent rich and poor school districts in the state.

Overall, Malloy's education plan takes us in the wrong direction. That said, if enacted, it might end the Connecticut paradox. We might end up with both low-quality average statewide performance and the largest achievement gap in the nation.

---

[i] **Tables describing the negative effect of teacher evaluations on student achievement among over 200,000 students in 64 countries (source: author's calculations based on data from the OECD PISA study)**

Table 1: The Negative Effect of using tests to judge teacher effectiveness on math achievement among over 200,000 students in 64 countries.

Effect of Purpose of the Assessments on Math Achievement

(legend: * p<0.05; ** p<0.01; *** p<0.001)

The Purpose of Assessments	Model 10a No Controls		Model 10b		Model 11 Full Model	
to compare with the national average	-7.43	***	-2.32	***	-1.43	**
to evaluate year to year progress	-8.64	***	-8.98	***	-9.40	***
to judge teacher effectiveness	-17.49	***	-15.74	***	-10.12	***
to inform parents of child's progress			**36.27**	***	**23.13**	***
to make decisions about promotion			-10.88	***	-1.24	*
for instruction			-4.05	***	-8.61	***
to improve the curriculum			**5.25**	***	**2.60**	***
to compare with other schools			-9.37	***	-6.21	***

Table 2: The Negative Effect of the increased use of standardized tests on math achievement among over 200,000 students in 64 countries.

Effect of Frequency of Assessments on Math Achievement

(legend: * p<0.05; ** p<0.01; *** p<0.001)

	Model 8a No Controls		Model 8b		Model 9: Full Model	
Frequency of Standardized Tests						
1-2 times a year	5.64	***	3.65	***	-0.52	
3-5 times a year	-9.37	***	-11.38	***	-7.98	***
once a month	-29.53	***	-29.43	***	-16.37	***
more than once a month	-28.50	***	-28.89	***	-18.19	***
Frequency of Teacher Judgemental Ratings						
1-2 times a year			21.75	***	16.46	***
3-5 times a year			9.13	***	2.67	***
once a month			2.23	*	-1.71	*
more than once a month			16.23	***	8.03	***
Frequency of Student Portfolios						
1-2 times a year			10.07	***	4.24	***
3-5 times a year			4.51	***	4.55	***
once a month			-6.87	***	-0.91	
more than once a month			-7.87	***	-0.41	